

創業企業の信用リスクモデル

尾木 研三^{†*} 内海 裕一[†] 枇々木 規雄[※]

Abstract

わが国では、80年代後半から企業数の減少が続いている。創業を増やすため、政府はさまざまな支援策を打ち出しており、この動きを受けて、銀行も創業企業への融資を積極化している。実績が増えるにつれて、創業企業の信用リスク計測が課題になってきた。中小企業の信用リスク計測には、主に決算書の数値から統計手法を用いて信用リスクを数値化する信用リスクモデルが使われる。ただ、これから創業する企業は決算書がないため、既存のモデルは使用できない。また、創業企業向けのモデルや非財務変数だけでリスク評価するモデルは、われわれの知る限り存在しない。

先行研究をみると、Gonçalves et al.(2014)は、1,430社のポルトガルの創業企業の3年間のパネルデータを使って、デフォルトの決定要因を分析している。鈴木(2012)は、2,897社のわが国の創業企業の5年間のパネルデータを用いて、廃業要因を分析している。これらの研究は、いずれも要因分析が目的であり、信用リスクモデルの構築や評価は行っていない。また、創業前の非財務変数だけではなく、創業後の財務変数も使用している。そこで、本研究では、日本政策金融公庫が保有する34,470社の創業企業の創業前の非財務変数だけを用いて、創業企業の信用リスクモデルを構築する。説明変数の選択にあたっては、デフォルト要因を「人的要因」「金融要因」「業種要因」の三つのカテゴリーに分けてロジット分析する。

分析の結果、人的要因として「開業の計画性」「起業経験年数」「年齢」などが有意になった。経営に必要な知識や体力が創業者に備わっているかどうかを評価していると考えられる。また、金融要因として、創業者の資産負債状況が有意になった。主に創業者個人の資金調達力を示している可能性がある。最後に、業種要因として、7つの業種グループのダミー変数が有意になった。デフォルト率の低い4つの業種グループは収益率が高いか競争が少ないという特徴がある一方、デフォルト率が高い3つの業種グループは、収益率が低いか競争が激しいという特徴があり、業界の経営環境を表していると思われる。分析で有意になった変数を用いてモデルを構築した結果、AR値は57%となり、実務でも利用可能であることが確認できた。

1 はじめに

わが国では、80年代後半から廃業企業数が創業企業数を上回るようになり、企業数の減少が続いている。創業を増やすため、創業の意識喚起や経営ノウハウに関するセミナーの開催、融資や保証制度の充実など、政府はさまざまな支援策を講じている。創業支援の機運が高まるなか、金融機関も創業企業に対する融資制度を創設して積極的に融資している。徐々に創業企業向けの融資件数や融資金額が増えるなかで、創業企業の信用リスクの計測が課題となっている。

最初に、本研究の対象とする創業企業の定義を確認する。公庫の新規開業資金は事業開始後おおむね7年以内が対象である。東京商工会議所も創業後5年未満を保証の対象としており、いずれも創業した後の企業を含めている。一方、「中小企業の新たな事業活動の促進に関する法律」では、「創業者」とは新たに事業を開始する具体的な計画を有するもの」としており、創業した後の企業は含めていない。このように、定義は必ずしも明確ではないが、本研究では、この法律の定義に従い、創業する前の企業を創業企業と定義する。したがって、創業を計画したけれども、何らかの事由によって創業できなかった企業も含ま

[†] (株)日本政策金融公庫国民生活事業本部リスク管理部, E-mail: ogi-k@jfc.go.jp (本稿で示されている内容は、筆者たちに属し、日本政策金融公庫としての見解をいかなる意味でも表さない。)

* 慶應義塾大学大学院理工学研究科

※ 慶應義塾大学理工学部管理工学科

れる。

中小企業の信用リスク計測には、主に決算書の数値から統計手法を用いて数値化する信用リスクモデルが使われる。モデルにはさまざまなタイプがあるが、ロジスティック回帰モデルが普及しており、様々な研究が行われている。国内では、AR¹値¹を用いてモデルの有効性を評価する研究(山下・川口・敦賀(2003)、柳澤他(2007)、三浦・山下・江口(2008)、枇々木・尾木・戸城(2011)、尾木・戸城・枇々木(2015))、財務変数に加えて非財務変数として、業歴の有効性を検証した研究(枇々木・尾木・戸城(2010))、マクロファクターの有効性を検証した研究(森平・岡崎(2009)、森平(2009)、枇々木・尾木・戸城(2012))、マーソンの1ファクターモデルとの整合性を示した研究(尾木・森平(2013))、ELの推計に関する研究(尾木・戸城・枇々木(2016))、変数選択とデータ量に関する研究(山下・川口(2003))などがある。国外でも、Ohlson(1980)が行ったロジットモデルを使ったデフォルト確率推定の研究に始まり、Altman and Sabato(2007)の研究など、数多くの研究がある。

どの先行研究も説明変数として財務変数を使用しているが、創業する前の企業には決算書がない。尾木・戸城・枇々木(2014)は、財務変数に加えて、非財務変数として業歴が有効であることを明らかにしたが、創業前の企業には業歴もない。以上のように、業歴のない創業企業を対象に、非財務変数のみで信用リスクを評価するモデルに関する研究は、筆者らの知る限り存在しない。

創業した企業のデフォルト要因に関する研究として、Gonçalves et al.(2014)は、2005年と2006年に創業したポルトガルの企業1,430社について、3年間の財務変数と非財務変数のパネルデータを用いてデフォルトの決定要因を分析している。

「金融資本」「人的資本」「産業活力」の三つのカテゴリーに分けて分析した結果、金融資本として、総負債対資本比率、資産対売上高比率、Ebitda対負債比率が有意になり、総負債対資本比率は、創業者個人の支援が大きな企業ほどデフォルト率が低いことを指摘している。また、人的資本として、学歴とマネジメント経験年数が有意になることを示した。産業活力については、業界の付加価値額の増加率、新規参入率(創業企業数÷既存企業数)などの変数を用いて分析しているが、いずれも非有意であることを述べている。

鈴木(2012)は、わが国で創業した企業の廃業要因を「起業家要因」「企業・環境要因」「戦略要因」の三つの視点からロジット分析している。日本政策金融公庫国民生活事業本部(以下、公庫という。)が融資した2006年に創業した企業12,190社に公庫の総合研究所がアンケート調査を行い、第1回(2006年12月末)のアンケートに回答した2,897社を対象に、2010年度まで毎年末アンケートを実施してパネル分析している(5回すべてのアンケートに回答した企業は783社である)。存続と廃業は、アンケートの自己申告のほか、訪問調査と公庫の債権管理情報を使って確認している。廃業を自発的廃業と非自発的廃業(リスク管理債権に分類されたもの)に分けて分析しており、非自発的廃業はデフォルトの定義に近いものとして参考になる。分析の結果、非自発的廃業は、起業家要因として、斯業経験年数、開業直前の職業(非正社員)、経営経験の有無、性別(女性が有意)、開業目的(やりがい重視か、収入重視か)が有意になっている。また、企業・環境要因として、開業時の従業員数、業種別廃業率、戦略要因として事業の新規性の高さ(自己評価)が有意にな

¹ AR 値の概要については付録を参照されたい。

ったことなどを示した。

いずれの研究も創業した企業のデフォルト要因や廃業要因について分析しているが、信用リスクモデルの構築と評価は行っていない。さらに、いずれも財務変数を用いて分析しているほか、鈴木（2012）は、信用リスク評価において重要な要因である金融面での分析が十分とはいえない。そこで、本研究では、公庫が保有する 34,470 社の創業する前の企業の非財務変数だけを用いて、創業企業の信用リスクモデルを構築する。説明変数の選択にあたっては、デフォルト要因を「人的要因」「金融要因」「業種要因」の三つのカテゴリーに分けて分析する。本研究と先行研究の違いをまとめると表 1 のとおりである。本研究の特徴は、創業する前の企業を対象としている点、それに伴い、非財務変数のみを用いている点、さらに、創業企業を対象にした先行研究に比べてデータ数が多いという点である。

表 1 本研究と先行研究の相違点

		Gonçalves et al.(2014)	鈴木(2012)	尾木ら(2014)	本研究
リスクモデルの構築		×	×	○	○
対象企業		創業した企業	創業した企業	業歴1年以上の企業	創業する前の企業
使用変数		財務・非財務変数	財務・非財務変数	財務・非財務変数	非財務変数
評価の視点	人的要因	○	○	○	○
	金融要因	○	×	○	○
	業種要因	○	○	○	○
データ数	max	1,430	2,897	1,089,362	34,470
	min		783		1,718
融資年度		2005-2006年度	2006年度	2004-2011年度	2011-2013年度

本研究の主な結果と貢献は以下のとおりである。

- (1)人的要因として「開業の計画性」「起業経験年数」「年齢」などが有意になった。事業に必要な知識や体力などが創業者に備わっているかどうかを評価していると考えられる。
- (2)金融要因として、創業者の資産負債状況が有意になった。創業者の資金調達力と計数観念の水準を示していると思われる。
- (3)業種要因として、付加価値率が高いか新規参入率が低い業種グループがプラス、付加価値が低い新規参入率が高い業種グループはマイナスで有意になった。付加価値の高さと競争環境がデフォルトに影響していると考えられる。
- (4)モデルの AR 値は 57% となり、実務でも利用可能であることが確認できた。

ただ、これらの分析に対して考察の記述が不十分な点もある。この理由は実務上の都合で説明変数を明示できず、分析結果に対する説明変数の符号条件や係数の大きさに関する議論ができなかったことである²。その一方で、先行研究にはない本研究の貢献を以下にまとめる。

- (1)わが国の創業企業向けの融資データを使って信用リスクモデルを構築し、実証分析を含めて評価を行った最初の論文である。
- (2)さらに、先行研究は、創業した企業の財務情報も使って分析しているのに対し、本研究は創業する前の企業を対象に非財務変数のみで企業評価を行った最初の論文である。

² 説明変数を明らかにすることは、信用リスクモデルを融資に利用する場合、学術的には問題なくても実務的には問題を引き起こす可能性がある。説明変数が明らかになれば、スコアが有利になるように申告するなど、偽装や粉飾を誘発する懸念があるからである。

(3)先行研究はロジット分析の変数の有意性のみを示しているのに対し、本研究では信用リスクモデルで重要な AR 値による評価を行い、実務でも利用可能な結果が得られた。

(4)先行研究で用いられたデータ数は限りがあり、結果の安定性に問題がある。それに対し、本研究では最大で 34,470 件のデータを用いており、信頼できる結果が得られている。

本論文の構成は以下のとおりである。第 2 章で分析の概要について述べ、第 3 章では、モデルに有効な説明変数を選択するためのロジット分析の結果とモデルの精度について検証する。第 4 章でまとめと今後の課題を述べる。

2 分析の概要

2.1 目的とフレームワーク

本分析の目的は、公庫が 2011 年度から 2013 年度の 3 年間に融資したこれから創業しようとする創業前の企業 34,470 社の非財務データを使って、創業企業のリスクを計測するための信用リスクモデルを構築し、実務で利用可能かどうかを検証することにある。

信用リスクモデルにはさまざまなタイプがある。なかでもロジスティック回帰モデル(ロジットモデル)は、最も一般的に用いられており、CRD (Credit Risk Database) 協会や RDB (日本リスク・データ・バンク(株))といった代表的な中小企業向けモデルでも採用されている。ロジットモデルには、複数の格付を直接推定する順序ロジットモデルと、デフォルトの有無を被説明変数としてデフォルト確率を推定する二項ロジットモデルとがある。

公庫は二項ロジットモデルを採用しており、算出されたデフォルト確率を概ね 0~100 点までのスコアに変換して使用している。本分析でも同じタイプのモデルを構築する。具体的なモデルの構築方法は 2.3 節で説明する。

本研究におけるデフォルトの定義は、 t 年度に融資した創業する前の企業のうち、 $t+1$ 年度に破綻懸念先以下に遷移した企業である。たとえば、2011 年度のデフォルト率とは、2011 年度に融資した企業のうち、2012 年度中にデフォルトした企業の割合である。ここで、分母には t 年度にデフォルトした企業、つまり、2011 年度中にデフォルトした企業は含んでいないことに注意してほしい。

モデル構築のポイントは変数選択である。変数次第でモデルの精度が左右され、有効な変数を一つ見つけただけで性能が大きく改善することもある。先行研究を参考に、デフォルトに影響を与える要因を、以下の三つのカテゴリーに分けて分析し、変数候補を抽出する。ヒト、モノ、カネは経営の三要素といわれている。このうち、モノについては、資格や技術などのソフトは創業者の能力に依存するし、店舗や設備などのハードは、創業前企業の場合は開業計画書を評価するしかない。したがって、モノの要素は人的要因に含めた。また、創業においては参入する業界の環境が大きく影響する。Gonçalves et al.(2014)、鈴木 (2012)も業種を一つのカテゴリーとして分析しており、業種要因を加えた。パラメータの推定はSAS/STAT®を用いる。

①人的要因：経営者の資質、業界経験、開業の計画性などを表す変数

②金融要因：創業者の資産負債状況などを表す変数

③業種要因：デフォルト率や企業数、業種の類似性などでまとめた任意の業種グループ

2.2 データの概要

2.2.1 使用変数

先行研究を参考に、カテゴリーごとの主な変数を示すと表 2 のとおりである。これから創業する企業は決算書がないため、変数はすべて非財務変数である。また、*印の変数は、データの都合上、加工作業や修正などが別途必要になる変数である。

表 2 主な使用変数

カテゴリー	主な変数
人的要因	斯業経験(現在の事業に関連する仕事の経験)年数* 創業前勤務先の勤務年数* 経営経験年数* FC加盟の有無* 転職回数* 教育水準* 創業計画の妥当性* 創業時の年齢 など
金融要因	創業時の現預金額* 不動産所有の有無 創業前の年収* 創業費用* 負債額 債務の履行状況 その他負債状況 など
業種要因	業種別デフォルト率 創業企業数 など

2.2.2 使用データ

使用データは表 3 のとおりである。公庫が 2011 年度から 2013 年度に融資したすべての創業前の企業 34,470 社のデータベース (DB1) を使用する。ただ、*印の変数については、34,470 社のうち業種と地域に比例して抽出したサンプル企業 1,718 社について、加工・修正作業を行ったデータベース (DB2) を使用する。

表 3 使用データ

	カテゴリー	融資年度	デフォルト年度	サンプル数
DB1	業種 金融資本(*印を除く) 人的資本(*印を除く)	2011-2013	2012-2014	34,470社
DB2	金融要因(*印) 人的要因資本(*印)	2011-2012	2012-2013	1,718社

2.2.3 サンプルバイアス

使用データにおいて考えられるサンプルバイアスは以下のとおりである。

- ①公庫を利用しなかった創業企業のデータは含まれない。たとえば、銀行や VC、自己資金だけで創業した企業が含まれていない。
- ②融資審査というフィルターを通しており、創業企業のなかで比較的良質な企業が対象となっている可能性がある。

③公庫の営業区域になっていない沖縄県の企業が含まれていない。

また、鈴木(2012)や尾木ら(2014)の研究も公庫の融資先のデータを用いており、結果が類似する可能性が高いことに注意する必要がある。

2.3 分析手順

2.2.2 で示したとおり、変数によって使用できるデータベースが異なる。そこで、表 4 のとおり、三つのモデルを構築する。まず、DB1 を用いて、カテゴリーごとにロジット分析を行う。分析で有意になった変数を使ってモデル 1 を構築する。次に、金融資本と人的資本の*印の変数について、DB2 を使ってロジット分析を行い、モデル 2 を構築する。最後に、モデル 1 とモデル 2 の各スコアを変数とするモデル 3 (最終モデル) を構築する³。

表 4 構築するモデル

	使用DB	使用変数	目的
モデル1	DB1	業種要因 金融要因(*印を除く) 人的要因(*印を除く)	カテゴリーごとに、有効な変数を抽出する
モデル2	DB2	金融要因(*印) 人的要因資本(*印)	*印について、有効な変数を抽出する
モデル3	DB2	モデル1のスコア モデル2のスコア	最終モデルの構築

具体的には次のとおりそれぞれのモデルを構築する。

(モデル 1) DB1 が保有する創業企業 i の J 個の人的要因に関する変数 $h_{1,i,j}(i = 1, \dots, N_1; j = 1, \dots, J)$ 、 K 個の金融要因に関する変数 $f_{1,i,k}(i = 1, \dots, N_1; k = 1, \dots, K)$ 、 M 個の業種要因に関する変数 $g_{i,m}(i = 1, \dots, N_1; m = 1, \dots, M)$ を使用して、ロジットモデルを構築し、最尤法によってパラメータ $\beta_{1,j}(j = 1, \dots, J)$ 、 $\gamma_{1,k}(k = 1, \dots, K)$ 、 $\delta_g(m = 1, \dots, M)$ を推定する。ここで、 $p_{u,i}$ は t 年に融資した創業企業 i が $t+1$ 年にデフォルトする確率、 u はモデル番号、 N_1 は DB1 の創業企業数を表す。 $z_{u,i}$ が大きいほどデフォルト確率は低くなる。

$$p_{1,i} = \frac{1}{1 + e^{z_{1,i}}}, z_{1,i} = \ln\left(\frac{1 - p_{1,i}}{p_{1,i}}\right) = \alpha_{1,0} + \sum_{j=1}^J \beta_{1,j} h_{1,i,j} + \sum_{k=1}^K \gamma_{1,k} f_{1,i,k} + \sum_{m=1}^M \delta_m g_{i,m} \quad (2.1)$$

(モデル 2) DB2 が保有する創業企業 i の X 個の人的要因に関する変数 $h_{2,i,x}(i = 1, \dots, N_2; x = 1, \dots, X)$ 、 Y 個の金融要因に関する変数 $f_{2,i,y}(i = 1, \dots, N_2; y = 1, \dots, Y)$ 、を使用して、ロジットモデルを構築し、最尤法によってパラメータ $\beta_{2,x}(x = 1, \dots, X)$ $\gamma_{2,y}(y = 1, \dots, Y)$ を推定する。ここで、 N_2 は DB2 の創業企業数を表す。

$$p_{2,i} = \frac{1}{1 + e^{z_{2,i}}}, z_{2,i} = \ln\left(\frac{1 - p_{2,i}}{p_{2,i}}\right) = \alpha_{2,0} + \sum_{x=1}^X \beta_{2,x} h_{2,i,x} + \sum_{y=1}^Y \gamma_{2,y} f_{2,i,y} \quad (2.2)$$

³ DB1 と DB2 の変数をすべて投入して推定する方が効率的であるが、DB2 のサンプル企業数は DB1 の約 5%に過ぎず、モデルの精度に与える影響が大きいと判断して別々に推計した。

(モデル 3) モデル 1 で算出した $z_{1,i}$ モデル 2 で算出した $z_{2,i}$ を変数として、DB2 を使ってロジットモデルを構築し、最尤法によってパラメータ B 、 C を推定する。

$$p_{3,i} = \frac{1}{1 + e^{z_{3,i}}}, z_{3,i} = \ln\left(\frac{1 - p_{3,i}}{p_{3,i}}\right) = A + Bz_{1,i} + Cz_{2,i} \quad (2.3)$$

それぞれのモデルにおいて推定されたパラメータを用いて計算された $z_{u,i}^*$ から企業 i の信用スコア $CS_{u,i}$ を計算する。

$$CS_{u,i} = \eta_0 + (\eta_1 - \eta_0) \left(\frac{z_{u,i}^* - Z_u(1\%)}{Z_u(99\%) - Z_u(1\%)} \right) \quad (2.4)$$

ここで、 $Z_u(1\%)$ 、 $Z_u(99\%)$ はモデル構築時のインサンプルデータにおける $z_{u,i}^*$ の 1 パーセント点、99 パーセント点を表す。これは信用スコアが $z_{u,i}^* = Z_u(1\%)$ ならば η_0 点、 $z_{u,i}^* = Z_u(99\%)$ ならば η_1 点となるように基準化している。本研究では、 $\eta_0 = 10$ 、 $\eta_1 = 90$ としている。 $z_{u,i}^*$ を直接用いても結果に影響を与えない。

3 分析結果

3.1 モデル 1 の構築

DB1 の 34,470 社のデータを用いて、「人的要因」「金融要因」「業種要因」の三つのカテゴリーごとにロジット分析を行い、変数候補を選択する。その後、それらの変数候補を用いてモデル 1 を構築する。具体的な手順は①～④のとおりである。

- ①「業種要因」について、業種別デフォルト率と企業数をもとに業種をグルーピングし、変数候補として有効な業種ダミーを抽出する。
- ②「金融要因」について、創業者の資産負債変数をもとにロジット分析し、有効な変数候補を抽出する。
- ③「人的要因」について、創業者の年齢を用いて有効な変数を抽出する。
- ④各カテゴリーの有効な変数候補を用いてモデル 1 を構築し、AR 値を算出する。

3.1.1 業種要因の変数選択

Gonçalves et al.(2014)は、①業界成長力、②参入率（創業企業数÷既存企業数）、③産業集中度（既存企業に対する創業企業の占有率）のいずれも非有意であることを示した。一方、鈴木（2012）は、廃業率の高い業種として「飲食店・宿泊業」「情報通信業」「小売業」を、低い業種として「医療・福祉」「個人向けサービス業」「運輸業」を示した。さらに、業種別廃業率が高い業種ほど非自発的廃業の廃業確率が有意に高まることを明らかにした。また、尾木ら(2014)は、業種によって業歴別デフォルト率の水準や形状が異なり、業界の経営問題や構造問題を反映していることを述べている。ポルトガルでは業種に関する変数がいずれも非有意であったものの、わが国では、業種別デフォルト率が業界環境を示す有力な変数候補になる可能性がある。

そこで、本項ではモデル構築に有効な業種ダミーを抽出するために、業界環境が類似した業種をグルーピングする。最初に、DB1 を使って小分類ベースでデフォルト率を算出する。そのあと、すべての小分類業種が平均デフォルト率近傍（2.0%±0.5%）にある製造業

と件数が少ない業種を除外する（大分類ベースで「農業、林業」「漁業」「鉱業、採石業、砂利採取業」「電気・ガス・熱供給・水道業」「金融・保険業」「公務」など）。残った 27,501 社の 87 業種についてデフォルト率を算出した結果を表 5 に示す。デフォルト率の水準は 0%～20%の間でさまざまである。創業企業数が 100 社以上の業種に限ってみても 0.0%～10.6%であり、業種によってデフォルト率に大きな差があることがわかる。市場の成長性や競争環境、収益性、参入障壁など、業界環境の違いが影響していると思われる。

表 5 創業企業の業種別デフォルト率

小分類	企業数	デフォルト率	小分類	企業数	デフォルト率	小分類	企業数	デフォルト率
1	71	4.2%	30	318	0.6%	59	115	1.7%
2	158	5.1%	31	32	0.0%	60	19	0.0%
3	115	2.6%	32	13	0.0%	61	71	4.2%
4	263	4.6%	33	200	1.5%	62	31	0.0%
5	22	0.0%	34	47	4.3%	63	126	3.2%
6	40	0.0%	35	107	3.7%	64	254	2.0%
7	64	3.1%	36	14	7.1%	65	525	4.0%
8	100	2.0%	37	23	0.0%	66	731	3.0%
9	207	0.0%	38	49	4.1%	67	31	3.2%
10	277	0.4%	39	34	2.9%	68	26	0.0%
11	70	0.0%	40	23	0.0%	69	30	3.3%
12	4,236	0.6%	41	39	0.0%	70	37	13.5%
13	664	1.1%	42	14	0.0%	71	40	2.5%
14	426	0.5%	43	151	2.0%	72	39	0.0%
15	944	0.4%	44	637	2.2%	73	68	4.4%
16	23	0.0%	45	136	1.5%	74	10	20.0%
17	2,100	0.7%	46	90	1.1%	75	468	2.1%
18	125	2.4%	47	15	0.0%	76	195	2.1%
19	627	1.4%	48	128	2.3%	77	18	0.0%
20	157	4.5%	49	13	0.0%	78	161	10.6%
21	580	3.6%	50	39	0.0%	79	262	3.1%
22	306	2.0%	51	17	11.8%	80	54	1.9%
23	851	3.5%	52	49	8.2%	81	21	0.0%
24	3,465	3.4%	53	20	0.0%	82	13	0.0%
25	570	2.1%	54	142	0.7%	83	55	1.8%
26	3,089	3.7%	55	22	4.5%	84	27	7.4%
27	25	4.0%	56	106	1.9%	85	1,593	0.3%
28	43	2.3%	57	20	5.0%	86	49	0.0%
29	67	4.5%	58	31	0.0%	87	218	0.9%

企業数が少ないと、1 社のデフォルトが業種別デフォルト率に与える影響が大きくなる。一定のサンプル数を確保するため、87 の業種について、デフォルト率の水準と業種の類似性をもとに、さらに 16 の業種にグルーピングを行った。この 16 の業種グループについて、ロジット分析をした結果を表 6 に示す。7 つの業種グループが有意になった。業種グループダミーだけでモデルを構築すると、AR 値は 35.4%となり、業種要因だけでもある程度の序列性が認められる。

業種要因の分析結果をもとに、p 値が 5%以上で有意になった 7 つの業種グループについてダミー変数を変数候補として選択する。ちなみに、7 つの業種グループの企業数は 24,324 社で、全体企業数 34,470 社に占める割合は 70.6%である。デフォルト率の高い業種グループは「宿泊業、飲食サービス業」「卸売業、小売業」など、低い業種グループは「医療、福祉①」「生活関連サービス業」「学術研究、専門・技術サービス業③」などで、鈴木 (2012) の指摘と同様の傾向がみられる。

表 6 業種グループ

業種グループ（大分類ベース）	デフォルト率	企業数	標準化回帰係数	p値
サービス業（他に分類されないもの）	高	607	▲ 0.08	<0.001
学術研究、専門・技術サービス業①	低	62	0.25	0.961
学術研究、専門・技術サービス業②	高	64	▲ 0.02	0.304
学術研究、専門・技術サービス業③	低	654	0.09	0.041
生活関連サービス業	低	4,900	0.16	<0.001
医療、福祉①	低	3,493	0.16	<0.001
医療、福祉②	低	627	0.00	0.888
医療、福祉③	高	125	▲ 0.02	0.423
宿泊業、飲食サービス業	高	9,043	▲ 0.20	<0.001
運輸業①	高	110	▲ 0.03	0.082
運輸業②	低	318	0.03	0.402
卸売業、小売業	高	3,767	▲ 0.13	<0.001
教育、学習支援業	高	637	▲ 0.03	0.181
建設業	高	1,064	▲ 0.03	0.188
情報通信業	高	170	▲ 0.01	0.488
不動産業	低	1,860	0.19	<0.001

デフォルト率の差がどのような業界環境の影響を受けているのかを調べるために、競争環境と収益性の観点から分析する。具体的には、平成 24 年経済センサスー活動調査（総務省統計局）のデータを用いて、7つの業種グループについて、参入率（＝平成 23 年以降に開設した事業所数÷平成 23 年の事業所数）と付加価値率（＝付加価値額÷売上高）を算出する。

結果を表 7 に示す。デフォルト率の低い業種をみると、付加価値率が 50%を上回っており、付加価値の高い業種が多い。不動産業は付加価値率が 50%を下回っているが、参入率が最も低く、比較的競争が緩やかであると考えられる。

一方、デフォルト率の高い業種はいずれも付加価値率が 50%を下回っている。宿泊業、飲食サービス業の付加価値率は 44.0%であり、低いレベルではないものの、参入率が最も高く、他の業種グループに比べて競争が激しい環境にあると思われる。

表 7 7つの業種グループの参入率と付加価値率

業種グループ（大分類ベース）	デフォルト率	参入率（％）	付加価値率（％）
宿泊業、飲食サービス業	高	4.5	44.0
サービス業（他に分類されないもの）	高	2.2	30.6
卸売業、小売業	高	2.0	10.4
生活関連サービス業	低	1.8	55.5
学術研究、専門・技術サービス業③	低	2.4	52.6
医療、福祉①	低	3.1	51.7
不動産業	低	0.5	31.2

資料：総務省統計局「平成 24 年経済センサスー活動調査」から作成

注：業種グループは任意のものであり、日本標準産業分類の業種分類（大分類）とは一致しないことに注意してほしい。

3.1.2 金融要因の変数選択

尾木ら(2014)は、小企業の信用リスク評価において、財務変数と並んで経営者の個人資産額が有力な変数であることを述べている。また、Gonçalves et al.(2014)は、経営者個人の金融面での支援が大きいほど、デフォルト確率が低いことを示した。先行研究の結果を踏まえ、創業者の資金調達力を示すものとして、資産負債状況に関する変数についてロジット分析する。

まず、創業者個人の資産負債状況を示す 85 個の変数について単変数ロジットを行い、p 値 5%水準で有意な変数を選択する。次に、相関マトリックスを作成して 0.5 以上の相関のある変数、クラスター分析で同じクラスターにある変数について取捨選択を行い、金融要因の変数を 15 個に絞り込んだ。

15 個の変数を用いてロジットモデルを構築した結果を表 8 に示す。ステップワイズによって 6 個の変数が選択された。5 つが負債状況を示す変数であり、符号条件がマイナスであることから、負債状況を示す変数の値が大きいほどスコアが低くなることがわかる。一方で、資産状況を示す変数は一つが選択された。符号条件はプラスであり、値が大きいほどスコアが高くなることがわかる。

表 8 金融要因のロジットモデル

変数名	推計値	標準化回帰係数	p値
定数項	3.94	—	<0.001
創業者の資産状況	0.56	0.22	<0.001
創業者の負債状況1	▲0.29	▲0.07	0.105
創業者の負債状況2	▲0.14	▲0.15	<0.001
創業者の負債状況3	▲0.63	▲0.07	<0.001
創業者の負債状況4	▲0.11	▲0.06	<0.001
創業者の負債状況5	▲0.52	▲0.16	<0.001

図 1 にデフォルト企業と非デフォルト企業のそれぞれのスコア別構成比を示す。非デフォルト企業の方が高いスコアの構成比が高くなっており一定の判別力があることが確認できる。AR 値は 38.7%と金融要因の変数だけでもある程度の序列性が確認できる。

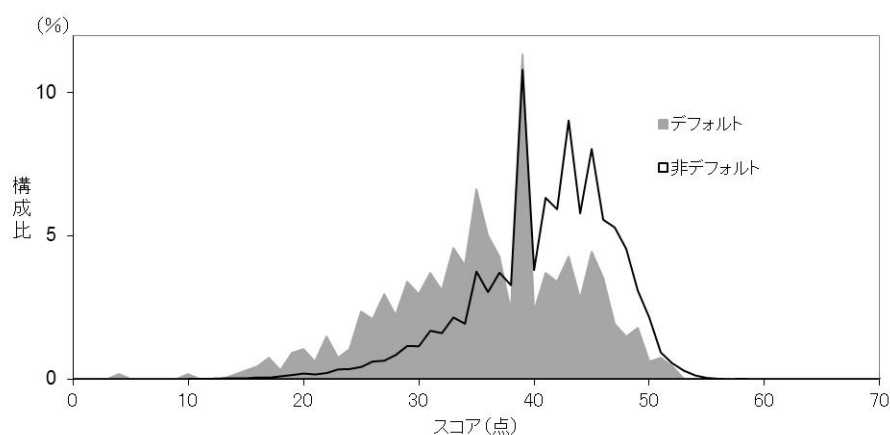


図 1 非デフォルト企業とデフォルト企業の金融要因スコア別構成比の分布

3.1.3 人的要因の変数選択

DB1 で使用できる人的要因は創業者の年齢のみであるが、創業時の年齢が創業後のパフォーマンスに与える影響は決して小さくないと考えられる。年齢と企業の経済的パフォーマンスとの関係について、玄田・高橋(2003)は、月商や付加価値が最大となる年齢は 40～42 歳であることを示している。また、創業者の年齢と廃業との関係について、鈴木(2012)は加齢に伴う知力と体力の低下が起業活動に与える影響について言及し、年齢の 1 次、2 次項を加えて推計した結果、全廃業は 44.6 歳でその確率が最も低くなることを示した。先行研究をみても、創業時の年齢が重要なファクターになる可能性が高いことがわかる。

図 2 にデフォルト企業と非デフォルト企業のそれぞれの年齢別構成比を示す。40 歳前後を過ぎるとデフォルト企業の構成比が非デフォルト企業の構成比を上回っている。先行研究では、40 歳代が起業家活動に必要な体力と知力のピークとしている。本分析でも 40 歳を過ぎると徐々にデフォルト企業の構成比が増えており、先行研究と同様の傾向が確認できる。

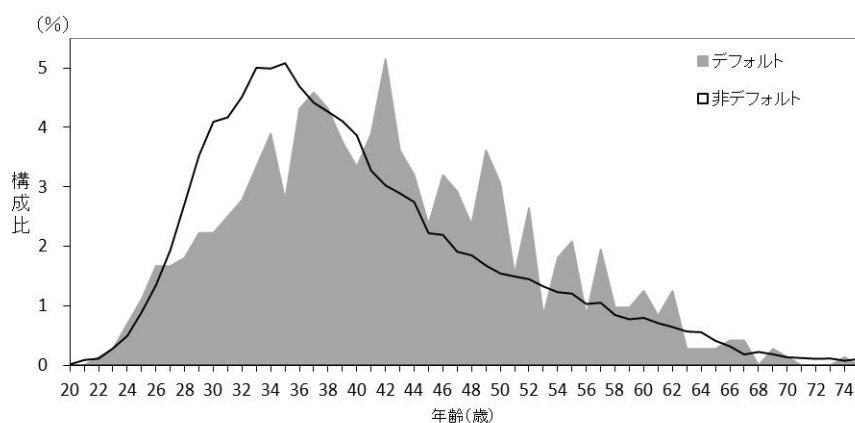


図 2 非デフォルト・デフォルト企業の年齢別構成比の分布

以上の検討を踏まえ、30 歳、35 歳、40 歳、45 歳、50 歳の 5 つのダミー変数について AR を計測した。結果を表 9 に示す。「創業時の年齢が 40 歳未満」のダミー変数の AR 値が 12.3%と、最も高くなった。

表 9 年齢ダミーの AR 値

ダミー変数	(単位: %)				
	30歳未満	35歳未満	40歳未満	45歳未満	50歳未満
AR値	2.1	9.7	12.3	8.7	4.4

3.1.4 モデル 1 の構築と評価

3.1.1～3.1.3 で抽出した変数を用いてモデル 1 を構築した。結果を表 10 に示す。変数は「学術研究、専門・技術サービス業③の業種ダミー」を除いてすべての変数が選択された。

図 3 のとおり、デフォルト企業と非デフォルト企業のスコア別構成比の分布がうまく分かれている。AR 値は 51.2%となり、モデル 1 だけでも実務で利用可能なレベルにあることが確認できる。

表 10 モデル 1 の標準化回帰係数

カテゴリー	変数名	推計値	標準化回帰係数	p値
	定数項	3.91	—	<0.001
業種ダミー	宿泊業、飲食サービス業	▲ 0.82	▲ 0.20	<0.001
	サービス業(他に分類されないもの)	▲ 0.71	▲ 0.06	<0.001
	卸売業、小売業	▲0.57	▲ 0.11	<0.001
	生活関連サービス業	0.65	0.13	<0.001
	医療、福祉①	0.84	0.14	<0.001
	不動産業	0.14	0.18	<0.001
金融資本	創業者の資産情報1	0.53	0.20	<0.001
	創業者の負債情報1	▲ 0.62	▲ 0.07	<0.001
	創業者の負債情報2	▲ 0.54	▲ 0.16	<0.001
	創業者の負債情報3	▲ 0.06	▲ 0.04	<0.001
	創業者の負債情報4	▲ 0.13	▲ 0.14	<0.001
	創業者の負債情報5	▲ 0.12	▲ 0.06	<0.001
人的資本	40歳未満ダミー	0.50	0.14	<0.001

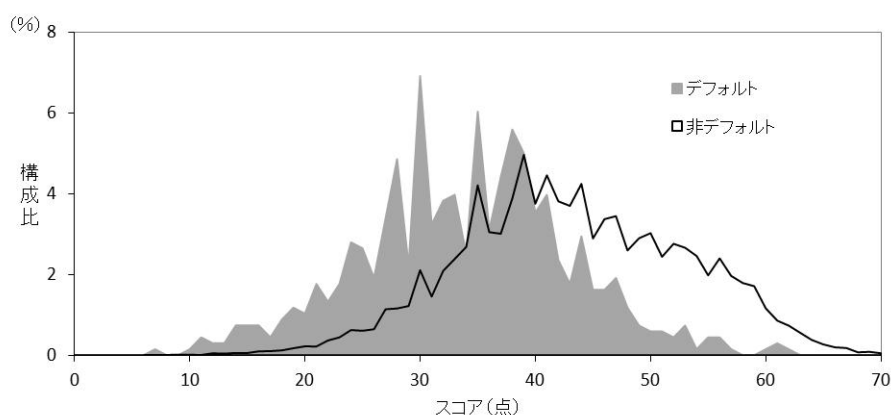


図 3 非デフォルト・デフォルト企業のスコア別構成比の分布(モデル 1)

3.2 モデル 2 の構築

「人的要因」「金融要因」のうち、*印のついた 10 個の変数について、DB2 を用いてロジットモデルを構築した。変数選択はステップワイズで行った。結果を表 11 に示す。人的要因として、斯業経験ダミー (5 年以内:1)、開業計画性の有無ダミー、金融要因として、創業者の手持ち資金額、の 3 つの変数が有意になった。Gonçalves et al.(2014)は、人的資本に関連する 5 つの説明変数を検証した結果、創業者の経営経験歴と教育水準が統計的に有意であったことを示した。鈴木(2012)は、非自発的失業に関しては、経営経験と教育水準のいずれも非有意であったことを述べている。本分析では、鈴木(2012)と同様に、経営経験や教育水準はいずれも有意にならなかった。

表 11 モデル 2 の標準化回帰係数

変数名	推計値	標準化回帰係数	p値
定数項	▲ 0.71	—	0.063
創業者の手持ち資金額	0.3	0.20	<0.001
斯業経験ダミー(5年以内: 1)	▲0.49	▲ 0.12	<0.001
開業計画性の有無ダミー(なし: 1)	▲0.50	▲ 0.11	0.006

斯業経験年数と廃業との関係について、鈴木（2012）は年数が長くなるほど確率が低くなることを示している。さらに、斯業経験年数と年ダミーとの交差項を作成して推計した結果、斯業経験の効果は創業後 3 年程度で消滅することを明らかにした。本分析では、1 年～10 年のダミーの AR 値を算出して検討した。結果を表 12 に示す。3 年以内の AR 値が 11.4%と最も高く、鈴木(2012)と整合的な結果となった。ただ、表 12 のとおり、4 年以内と 5 年以内はほぼ同水準で、6 年以内以上から徐々に低下している。そこで、本分析では、5 年以内ダミーを採用することにした。

表 12 斯業経験年数ダミーの AR 値

(単位: %)										
ダミー変数	1年以内	2年以内	3年以内	4年以内	5年以内	6年以内	7年以内	8年以内	9年以内	10年以内
AR値	8.4	10.4	11.4	10.9	11.1	9.0	6.6	6.9	8.1	5.9

創業計画の妥当性の有無は、審査担当者の主観に基づく情報であるが有意になった。また、金融要因として、自己資金額に親や配偶者、親戚などからの支援金額や借入金のうちで使途が確定していない余裕資金額などを加えた手持ち資金額が選択された。Gonçalves et al.(2014)の分析でも、創業の早い段階で経営者個人の金融面での支援が大きいほど、創業企業のデフォルト確率が低いという結果が出ており、先行研究とも整合的な結果となっている。

図 4 に、モデル 2 の非デフォルト・デフォルト企業のスコア別構成比の分布を示す。モデル 1 に比べて判別力が低いことがわかる。AR 値は 27.9%となった。

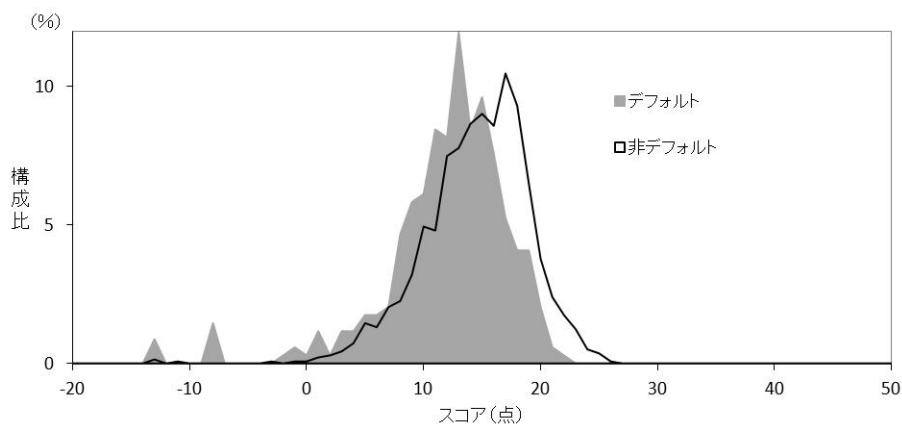


図 4 非デフォルト・デフォルト企業のスコア別構成比の分布（モデル 2）

3.3 モデル 3（最終モデル）の構築

DB2（1,718 社）のデータを用いて、モデル 1 のスコアとモデル 2 のスコアを算出し、それぞれのスコアを変数としてモデル 3（最終モデル）を構築した。結果を表 13 に示す。標準化回帰係数をみると、モデル 1 のスコアが 0.59、モデル 2 のスコアが 0.22 となった。

表 12 モデル 3 の標準化回帰係数
(修正疑似 $R^2=0.172$)

変数名	推計値	標準化回帰係数	p値
定数項	▲ 0.12	—	<0.001
モデル1のスコア	0.98	0.59	<0.001
モデル2のスコア	0.85	0.22	<0.001

図 5 にモデル 3 の非デフォルト・デフォルト企業のスコア別構成比の分布を示す。AR は 57.1% となり、実務でも十分活用できる水準にあることが確認できる。

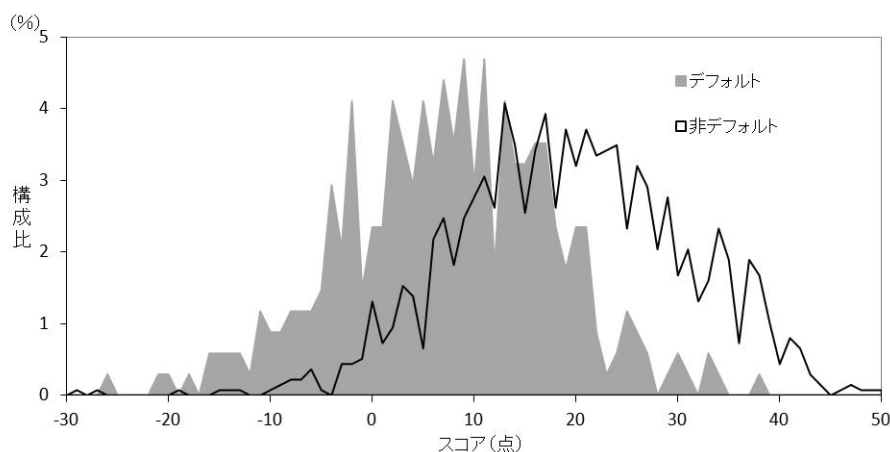


図 5 非デフォルト・デフォルト企業のスコア別構成比の分布（モデル 3）

4 まとめ

本研究では、34,470 社の創業企業のデータベースを用いて、信用リスクモデルを構築し、AR 値による評価を行うことによって、その特徴と有用性を明らかにした。著者たちの知る限りにおいて創業企業の信用リスクモデルに関する先行研究は存在しておらず、本研究の内容はきわめてオリジナルな成果と言えるだろう。

本研究と主な先行研究との比較をまとめたものを表 13 に示す。AR 値の水準は、尾木ら(2014)の小企業向けモデルの 42.7% よりも高い 57.1% と、実務でも利用可能であることがわかった。

変数選択にあたっては、「人的要因」「金融要因」「業種要因」の三つの視点から分析を行った。分析の結果、人的要因として、年齢、起業経験、創業計画が有意になり、鈴木(2012)とほぼ同様の結果となった。経営者として必要な知識や体力などが備わっているかどうかは創業の成否を分ける重要なポイントであり、その点を評価していると考えられる。

また、金融要因として、創業者個人の資産負債状況と手持ち資金額が有意になった。Gonçalves et al.(2014)や尾木ら(2014)も経営者の資金力はデフォルトと相関高いことを示している。所有と経営が一体となっている小企業の場合、経営者個人の資産や資金調達力は経営上の重要な要素であり、先行研究や現場の経験則とも整合している。

最後に、業種要因として、7つの業種グループのダミー変数が有意になった。デフォルト率の低い4つの業種グループは収益率が高いか競争が少ないという特徴がある一方、デフォルト率が高い3つの業種グループは、収益率が低い競争が激しいという特徴があり、妥当な結果が得られた。鈴木(2012)の結果と同様の傾向が確認できる。

表 13 本研究と先行研究の分析結果の比較

		Gonçalves et al.(2014)	鈴木(2012)	尾木ら(2014)	本研究
リスクモデルの構築		×	×	○(AR値:42.7%)	○(AR値:57.1%)
対象企業		創業した企業	創業した企業	業歴2年以上の企業	創業する前の企業
使用変数		財務・非財務変数	財務・非財務変数	財務・非財務変数	非財務変数
人的要因	年齢	—	加齢に伴う知力と体力の低下が起業活動に与える影響について言及し、全廃業は44.6歳でその確率が最も低いことがわかった	—	30歳、35歳、40歳、45歳、50歳の5つのダミー変数についてARを計測した結果、「創業時の年齢が40歳未満」のダミー変数が最も高いことがわかった
	経営経験	有意	非有意	業歴が長いほどデフォルト率が低いことがわかった	非有意
	教育水準	有意	非有意	—	非有意
	起業経験	非有意	有意(3年)	—	有意(5年)
	創業計画	非有意	有意 「FC加盟」 「新規性の有無」	—	有意 「創業計画の妥当性」
金融要因	資産負債状況	経営者個人の金融面での支援が大きいほど、デフォルト確率が低い	—	経営者の個人資産額が有力な変数である	6個の変数が有意 5つが負債情報 1つが資産情報
	手持ち資金額		—		有意
業種要因		①業界成長力 ②参入率 ③産業集中度 は非有意	業種を取り巻く環境が廃業に影響していることを指摘 廃業率の高い業種 「飲食店・宿泊業」 「情報通信業」 「小売業」 廃業率の低い業種 「医療・福祉」 「個人向けサービス業」 「運輸業」	業種によって業歴別デフォルト率の水準や形状が異なり、業界の経営問題や構造問題を反映していると指摘	業種ダミーが有意であることを示したデフォルト率の高い業種グループ 「宿泊業、飲食サービス業」 「卸売業、小売業」など デフォルト率の低い業種グループ 「医療、福祉①」 「生活関連サービス業」 「学術研究、専門・技術サービス業③」

実務で利用するために、具体的な説明変数を示すことはできなかったが、先行研究と整合的で、そのうえ現場担当者にも納得感のある「人的要因」「金融要因」「業種要因」に関連する説明変数でモデルは構築されており、創業企業の特徴も表すことができた。本研究で構築したモデルを創業したばかりの企業(業歴の浅い企業)に応用することができれば、小企業向けモデルのAR値の向上に寄与する可能性もある。

今後の課題として、データの蓄積を待って以下の2点について研究を進める予定である。

①アウトオブサンプルテストを行い、モデルの頑健性を確認したい。

②時系列での分析を行いたい。

本研究の結果は、創業企業においても信用リスクモデルは有用であることを示しており、公庫のみならず、他の金融機関に多少なりとも参考になれば幸いである。

付録 AR 値(Accuracy Ratio)

信用リスクモデルの精度を評価する指標はいくつかあるが、デフォルト確率（PD）を推計するモデルの評価は AR 値で行われることが一般的である。AR 値はモデルから算出される PD の高い企業ほど、実際のデフォルトが生じている状況にあるのかどうかを評価する順序性尺度である。具体的には、サンプル企業を PD の高い順番に並べたあと、全企業数に対する累積企業の割合を X 軸にとる。次に、全デフォルト企業数に対する累積デフォルト企業の割合を Y 軸にとってプロットする。プロットによって示される曲線は CAP 曲線（Cumulative Accuracy Profiles）と呼ばれている。

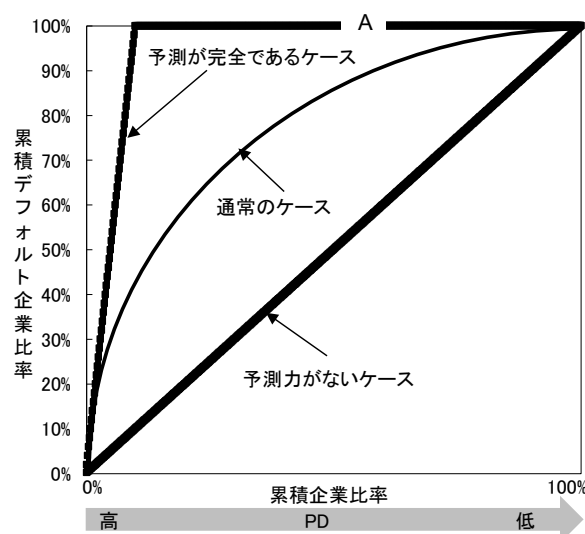


図 1 AR 値

図 1 の 45 度線に重なっている線は、PD の大きさと実際のデフォルトが全く関係なく生じており、モデルに予測力がないケースを示している。一方、線 A は、PD の高い順番に実際のデフォルトが生じたケースであり、モデルの予測が完全であるケースを表している。通常のケースは、45 度線と線 A の間に CAP 曲線が描かれる。したがって、45 度線と線 A で囲まれた面積（完全に予測したケース）を 1 とすると、AR 値は、線 A と 45 度線で囲まれた面積に対する CAP 曲線と 45 度線で囲まれた面積の比率であり、

$$AR値 = \frac{\text{CAP曲線と45度線で囲まれた面積}}{\text{線Aと45度線で囲まれた面積}}$$

で表される。モデルの精度が高ければ CAP 曲線は線 A に近い線を描き、精度が低ければ 45 度線に近い線を描くことになる。このように、AR 値は 0 から 1 の値をとり、1 に近いほど精度が高いという直観的なわかりやすさがあるうえ、格付付与においてはデフォルト確率の値の一致精度よりも、企業の信用リスクの相対的な大きさ（リスクの大きさの順序性）が重視されるという実務での有用性から広く普及している。

<参考文献>

- [1]尾木研三,戸城正浩,枇々木規雄(2016)「小企業の EL 推計における業歴の有効性」『ファイナンスにおける数値計算手法の新展開 (ジャフィー・ジャーナル:金融工学と市場計量分析)』,朝倉書店,2016 年 3 月,pp.156-178.
- [2]尾木研三,戸城正浩,枇々木規雄(2015)「小企業向け保全別回収率モデルの構築と実証分析」『ファイナンスとデータ解析(ジャフィー・ジャーナル:金融工学と市場計量分析)』朝倉書店,2015 年 3 月,pp.168-201.
- [3]尾木研三,戸城正浩,枇々木規雄(2016),「小規模企業向け信用スコアリングモデルにおける業歴の頑健性と経営者の個人資産との関係性」, Transactions of the Operations Research Society of Japan, Vol.59,掲載予定.
- [4]尾木研三,森平爽一郎(2013)「中小企業のデフォルト率に影響を与えるマクロ経済要因ー1ファクターモデルを用いたアプローチ」日本政策金融公庫論集第20号,2013年8月, pp.71-89.
- [5]玄田有史,高橋陽子(2003)「自己雇用の現在と可能性」,「調査季報」第 64 号,国民生活金融公庫総合研究所,pp.1-27.
- [6]鈴木正明(2012)「新規開業企業の軌跡 パネルデータにみる業績,資源,意識の変化」,勁草書房.
- [7]東京商工会議所 HP, <https://www.tokyo-cci.or.jp/shikin/wing/>
- [8]中小企業庁 HP,中小企業の新たな事業活動の促進に関する法律逐条解説(最終更新日平成 18 年 7 月 4 日),http://www.chusho.meti.go.jp/keiei/shinpou/chikujou_kaisetu/
- [9]日本政策金融公庫 HP,
https://www.ifc.go.jp/n/finance/search/01_sinkikaigyuu_m.html
- [10]枇々木規雄,尾木研三,戸城正浩(2010)「小企業向けスコアリングモデルにおける業歴の有効性」津田博史,中妻照雄,山田雄二編『定量的信用リスク評価とその応用(ジャフィー・ジャーナル:金融工学と市場計量分析)』朝倉書店,2010 年 3 月,pp.83-116.
- [11]枇々木規雄,尾木研三,戸城正浩(2012)「信用スコアリングモデルにおけるマクロファクターの導入と推定デフォルト確率の一致精度の改善効果」 Transactions of the Operations Research Society of Japan, Vol.55 (2012 年 12 月),pp.42-65.
- [12]枇々木規雄,尾木研三,戸城正浩(2011)「教育ローンの信用スコアリングモデル」津田博史,中妻照雄,山田雄二編『バリュエーション(ジャフィー・ジャーナル:金融工学と市場計量分析)』朝倉書店,2011 年 4 月,pp.136-165.
- [13]三浦翔,山下智志,江口真透(2008)「信用リスクスコアリングにおける AUC と AR 値の最大化法」,金融庁金融研究センター,20 年度ディスカッションペーパー.
- [14]森平爽一郎(2009)『信用リスクモデリングー測定と管理ー』,朝倉書店.
- [15]森平爽一郎,岡崎貫治(2009)「マクロ経済変数を考慮したデフォルト確率の期間構造推定」,早稲田大学大学院ファイナンス総合研究所ワーキングペーパーシリーズ.
- [16]柳澤健太郎,下田啓,岡田絵理,清水信宏,野口雅之(2007)「RDB データベースにおける信用リスクモデルの説明力の年度間推移に関する分析」,日本金融・証券計量・工学学会 2007 年夏季大会予稿集, 249-263.

- [17] 山下智志,川口昇,敦賀智裕(2003)「信用リスクモデルの評価方法に関する考察と比較」,金融庁金融研究センター,15年度ディスカッションペーパー.
- [18] 山下智志,川口昇(2003)「大規模データベースを用いた信用リスク計測の問題点と対策(変数選択とデータ量の関係)」,金融庁金融研究センター,14年度ディスカッションペーパー.
- [19]E.I.Altman and G.Sabato(2007),"Modelling Credit Risk for SMEs:Evidence from the U.S. Market",*ABACUS*,Vol.43,No.3,pp.332-357.
- [20]J.A.Ohlsion (1980),"Financial ratios and the probabilistic prediction of bankruptcy",*Journal of Accounting Research*,18(1),109-131.
- [21]Gonçalves et al.(2014),"The Determinants of Credit Default on Start-Up Firms.Econometric Modeling using Financial Capital, Human Capital and Industry Dynamics Variables",*FEP Working Papers*,n.534.